Water Research Vol. 15, pp. 1157 to 1171, 1981
Printed in Great Britain. 

0043-1354/81/101157-15$02.00/0


# A SYSTEMS MODEL OF STREAMFLOW AND WATER QUALITY IN THE BEDFORD OUSE RIVER SYSTEM—II. WATER QUALITY MODELLING

PAUL WHITEHEAD[1], BRUCE BECK[2] and ENDA O'CONNELL[1]

[1]Institute of Hydrology, Wallingford, England and
[2]International Institute for Applied Systems Analysis, Laxenburg, Austria

(*Received October* 1980)

**Abstract**—This paper, the second of a two part description of the modelling activities associated with the Bedford Ouse Study, concentrates on the theme of water quality modelling. In the first part of the paper streamflow models of a dynamic-stochastic type were developed and this approach has been extended during the water quality modelling studies. Recursive estimation techniques such as the extended Kalman filter have been applied to daily water quality data to determine models for dissolved oxygen, biochemical oxygen demand, chloride and nitrate. The models have been developed specifically to investigate design and operational management problems and have been used to assess the impact of effluent from the new city of Milton Keynes on river water quality. Monte Carlo simulation techniques are used to investigate the distributions of water quality at a downstream abstraction site and an application of the models linked to a water quality monitoring and telementry system is considered.

## INTRODUCTION

In part 1 of this paper (Whitehead *et al.*, 1979), a streamflow model of the Bedford Ouse river system has been presented. The approach discussed considers the dynamic and stochastic behaviour of river systems and characterises channel flow and rainfall-runoff behaviour using relatively simple low order models. The structure and complexity of these models is consistent with the objectives of the Bedford Ouse study which were to investigate design and operational management problems in the Bedford Ouse river basin. In this second part of the paper the water quality models developed for the Bedford Ouse river system are presented and an approach is considered for the design and operation of water resource systems from the viewpoint of water quality.

In addition to being major sources of water supply, rivers are used as the principal disposal pathways for waste materials. As discussed by the Director of the National Water Council (Scott, 1979) "The variety of pollutants... generated by a highly industrial society appears to grow continuously" and "the problems of water quality are now more difficult and demanding than water quantity". While in general, average water quality in the U.K. has tended to improve, in certain respects there have been grounds for concern. For example, some Water Authorities have been observing progressively increasing levels of nitrates in surface and groundwater systems (Onstad & Blake, 1980; Thompson, 1979). The mechanisms governing these increases are not wholly understood and yet strategies for the management of nitrate levels are now required in order to meet water quality standards. In particular, nitrate levels in the Bedford Ouse River at Bedford frequently exceed the WHO limit of 11.3 mg N $l^{-1}$ and it is necessary to blend the abstracted river water with low nitrate reservoir water and groundwater. The observation that certain acceptable quality limits are exceeded from time to time indicates, that stream quality should not be judged only in terms of, say, yearly average indices. Transient, intermittent deterioration of quality is also important, and may be of growing concern for the future (Beck, 1979a). The problems created by the variability of water quality are recognized widely, as is evidenced by the setting of percentile limits in EEC environmental directives and National Water Council guidelines.

Given the increasing levels of pollutants such as nitrates two questions may be posed: why are they increasing? and given that they are increasing, what can be done about them? In order to assess the mechanisms governing pollution, it is necessary to develop *descriptive* models which incorporate the detailed complex interactions between water quality variables, biological compartments and the physical environment. However, in order to provide management strategies a range of models are required which may be termed *prescriptive* and which are chosen carefully to suit the nature of the problem. In the case of the Bedford Ouse, water quality will be further exacerbated by the development of the new city of Milton Keynes (see Fig. 1). In this paper water quality models are identified for the critical stretch of the Bedford Ouse between the discharge point for Milton Keynes effluent and the Bedford Water Division abstraction plant, some 55 km downstream. In order to assess the likely impact of effluent on the river system and the operational management problems that may thereby arise, dynamic stochastic models of water quality

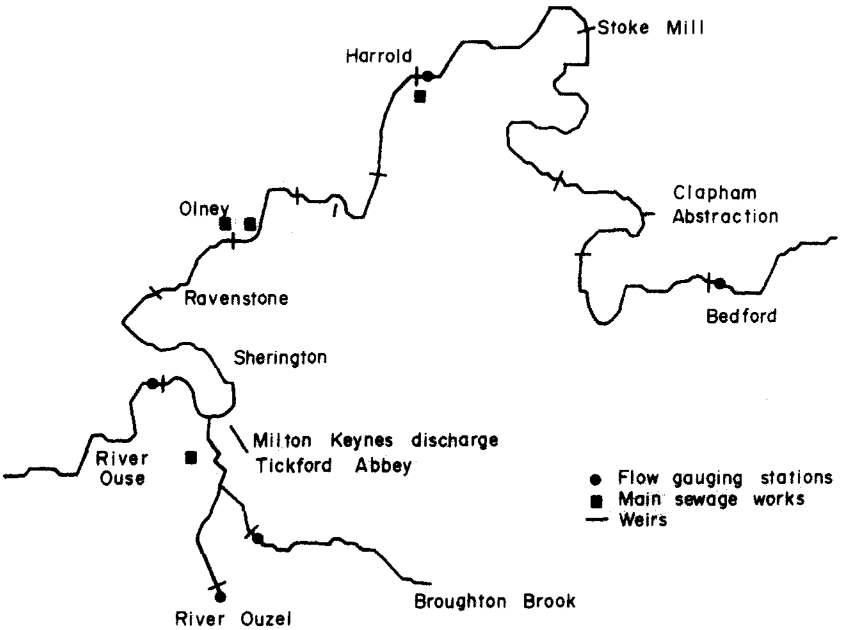


Fig. 1. Map of Bedford Ouse river system.

have been developed. Here such models are discussed together with the techniques required for their development and finally some applications of the models are presented.

## A HIERARCHY OF WATER QUALITY MODELS

There has been a tendency in recent years to categorize water quality models as either planning or operational management aids. However, such a breakdown is not strictly correct since planning models provide the "steady state" or annual average water quality conditions (Fawcett, 1975; Knowles & Wakeford, 1978; Casapieri *et al.*, 1978) which generally yields insufficient information for the *detailed* design of a water resource system. Moreover, the steady-state modelling approach has been questioned on the following technical grounds: the probability distribution of downstream water quality, which currently underlies the National Water Council's guidelines for the setting of consent conditions for effluent discharges, depends on short-term interactions between the upstream flow and quality and these interactions are not considered by a steady-state model (Warn, 1980). Thus the usefulness of the steady state modelling approach can be somewhat limited during the design phase and a dynamic modelling approach is required to account for the flow and water quality interactions.

The planning, design and operational phases of water resources management are interdependent since the optimality of any development plan will depend on the medium-term management policies adopted, and also on the short-term operational control of the water resource system. Clearly it would be computationally impossible to implement a single overall model which would, on the one hand, allow the evaluation of a large number of alternatives at the planning stage, and still maintain the level of detail required to represent the short-term behaviour of each system. This has led to the consideration of a *hierarchy* of models which can be identified with the planning, design and operational phases of water resources management (Jamieson, 1978).

### *Planning models*

Water quality planning models are characterized by their ability to evaluate a large number of potential investment programmes. Steady-state models, which, as discussed previously, may require gross assumptions about the system behaviour, are used in combination with an optimisation technique to evaluate alternative strategies for water resource development. Mixed-integer programming has been used for water quantity models (O'Neill, 1972) and dynamic programming techniques have been applied in the Trent Study (Newsome *et al.*, 1971). The Trent Mathematical Model is a well-known example of a water quality planning model.

### *Design models*

Following identification of a small number of near-optimal least cost solutions at the planning stage, these potential designs require considerably more detailed analysis to enable the optimal system to be selected and designed. The type of model required at this design stage must accommodate the natural flow and water quality fluctuations, typically on a daily basis, so that estimates can be made of the probability that certain water quality levels will arise at specified points within the proposed water resource system. This is precisely the information that is required for setting consent conditions for effluent discharge. The orientation in the design phase is therefore towards a dynamic, stochastic approach to water quality modelling.

Research into this class of water quality models has been undertaken by Page & Warn (1973) for evaluating the effect of effluent on water quality in a pumped storage system and the models developed in the Bedfored Ouse Study have likewise been applied to assess the impact of effluent on river water quality (Whitehead & Young, 1979) using a Monte Carlo simulation approach.

*Operational models*

The purpose of operational models is to assist in the efficient short term (day-to-day) management of water resource systems. The operational model uses data relating to the recent past and present state of the system in order to produce flow and water quality forecasts on which short term management decisions can be based. A typical application might be the protection of an abstraction point on a river system or the control of a multi-purpose reservoir system. The Bedford Ouse Study dynamic models are suitable for short term operational management and their application is considered later in this paper.

## DYNAMIC-STOCHASTIC WATER QUALITY MODELS

*Model structure*

Dynamic water quality models have been discussed in detail by Thomman (1972) and Rinaldi *et al.* (1979); a recent survey of the literature on applications of system identification and parameter estimation techniques in water quality modelling is given in Beck (1980).

If a model is to be useful for the purposes of detailed design and operational management, it should possess the following properties:

(i) It should be a truly dynamic model, being capable of accepting time-varying input (upstream) functions of water quality which are used to compute time-varying output (downstream) responses.

(ii) The model should be as simple as possible yet consistent with the ability to characterise adequately the important dynamic aspects of the system behaviour.

(iii) It should provide a reasonable mathematical approximation of the physicochemical changes occurring in the river system and should be calibrated against real data collected from the river at a sufficiently high frequency and for a sufficiently long period of time.

(iv) It should account for the inevitable errors associated with laboratory analysis and sampling, and account for the uncertainty associated with imprecise knowledge of the pertinent physical, chemical and biological mechanisms.

As indicated in Fig. 2, the structure of the Bedford Ouse model is to link the water quality model with the hydrological model such that interactions are incorporated directly. The water quality models developed in the Bedford Ouse Study are based on a transportation delay/continuously stirred tank reactor (CSTR) idealisation of a river (see Fig. 3). The mathematical formulation of this model is in terms of lumped parameter, ordinary differential equations and draws upon standard elements of chemical engineering reactor analysis, e.g. Himmelblaú & Bischoff (1968). As indicated in part 1 of this paper (Whitehead *et al.*, 1979) this idealization can be shown to approximate the analytical properties of the distributed-parameter, partial differential equation representations of advection-dispersion mass transport (see Rinaldi *et al.*, 1979); it can also be shown to approximate the experimentally observed transport and dispersion mechanisms (Whitehead, 1980a). The form of the model has been verified against water quality data in several studies (Thomann, 1972; Beck & Young, 1975).

The prinicpal advantages of this model over the corresponding partial differential equation descriptions are:

(a) the simplified computation required to solve the equations in the case of lumped parameter differential equations;

(b) the availability of statistically efficient algorithms for model identification and parameter estimation that can be readily applied to equations of the lumped-parameter form;

(c) the availability of extensive control system methods which are best suited to an ordinary differential equation model and which may be used for management purposes.

The mathematical form of the model is derived from a component mass balance across a reach of a river. For the CSTR shown in Fig. 3.

$$\frac{d\mathbf{x}(t)}{dt} = \frac{Q(t)}{V}\tilde{\mathbf{u}}(t) - \frac{Q(t)}{V}\mathbf{x}(t) + \mathbf{S}(t) + \boldsymbol{\zeta}(t) \qquad (1)$$

while for the transportation delay,

$$\tilde{\mathbf{u}}(t) = \mathbf{u}(t - \tau(t)) \qquad (2)$$

where

- $\mathbf{u}(t)$ is the vector of input, upstream component concentrations ($mg\,l^{-1}$)
- $\tilde{\mathbf{u}}(t)$ is the vector of (hypothetical) time-delayed input, upstream component concentrations ($mg\,l^{-1}$)
- $\tau(t)$ is the magnitude of the transportation delay element (day)
- $\mathbf{x}(t)$ is the vector of output, downstream component concentrations ($mg\,l^{-1}$)
- $\mathbf{S}(t)$ is the vector of component source and sink terms ($mg\,l^{-1}\,day^{-1}$)
- $\zeta(t)$ is the vector of chance random disturbances affecting the system ($mg\,l^{-1}\,day^{-1}$)
- $Q(t)$ is the stream discharge ($m^3\,day^{-1}$)

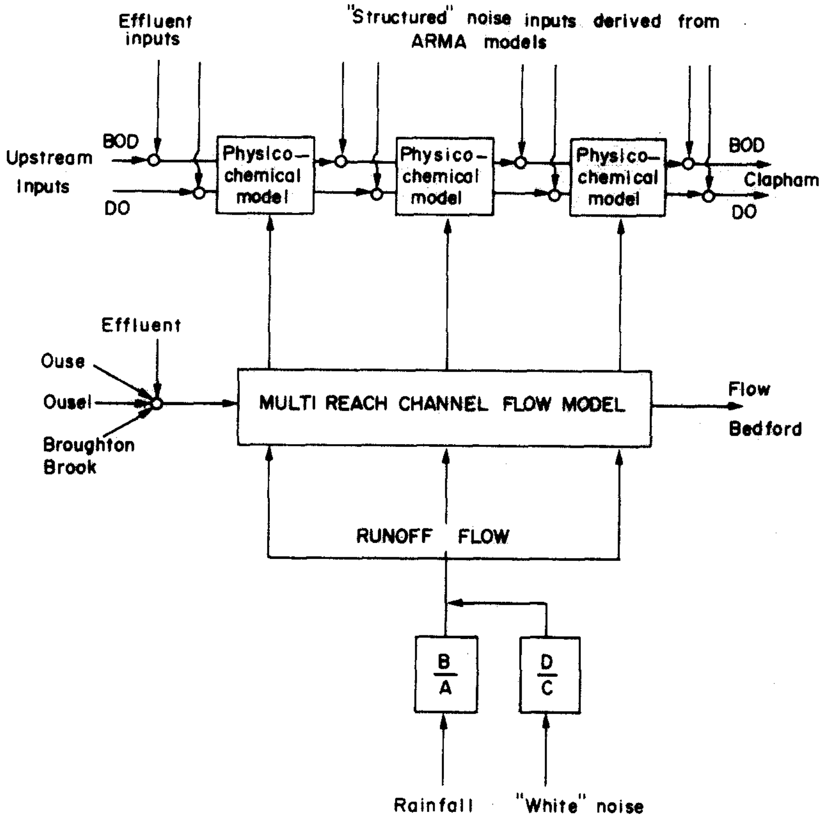


Fig. 2. Integrated model of flow and water quality.

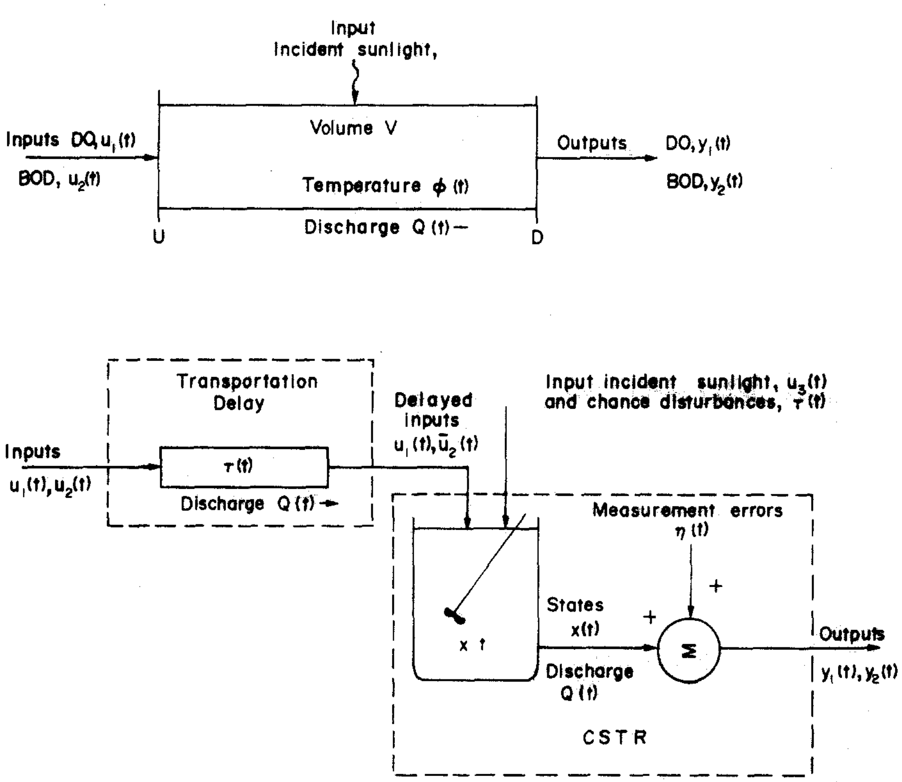


Fig. 3. Idealized river reach represented by transportation delay and CSTR model.

$V$ is the reach volume ($m^3$)
$t$ is the independent variable of time.

The errors associated with the laboratory analysis and sampling are included in the observation equation:

$$\mathbf{y}(t) = \mathbf{x}(t) + \mathbf{n}(t) \tag{3}$$

where

$\mathbf{v}(t)$ is the vector of observed (measured) downstream component concentrations ($mg\,l^{-1}$)

and

$\mathbf{n}(t)$ is a vector of the chance measurement errors ($mg\,l^{-1}$).

*Alternative discrete-time (grey box) models of water quality*

An alternative approach to modelling water quality is to use discrete-time state equations of the form:

$$\mathbf{x}_k = A\mathbf{x}_{k-1} + B\mathbf{u}_{k-1} \tag{4}$$

where $\mathbf{x}$ and $\mathbf{u}$ represent respectively the states (e.g. downstream BOD and DO concentrations and the inputs (e.g. upstream BOD and DO concentrations) while $A$ and $B$ are matrices of unknown parameters that are to be estimated from the field data.

The form of equation (4), when combined with the output observations of equation (3), leads to a multiple-input/multiple-output (MIMO) extension of the single-input/single-output (SISO) model discussed in Part I of this paper in relation to rainfall-runoff modelling. Such a MIMO model can be obtained by transforming the continuous-time, differential equation representation of equation (1) into the discrete-time representation of equation (4). For a linear description this can easily be achieved with standard techniques for the integration of vector differential equations (see, for example, Dorf, 1965). However, such a transformation will not normally lead to convenient expressions in which the parameters of the original (i.e. continuous-time) model description can be easily estimated. On the other hand, if the derived model of the system were specified directly in the form of equation (4), the inconvenience of the transformation is circumvented but the more immediate "physical" interpretation of the parameters in $A$ and $B$ is thereby sacrificed.

From the point of view of data analysis and system identification, there may be significant advantages in turning to the description of equation (4) as a model for water quality variations. These advantages, and an associated method of parameter estimation are discussed in detail in Young & Whitehead (1977). Since this alternative model is essentially a structured input/output representation, and thus might be called a grey box model, the problem of MIMO model identification is best approached by utilizing as much *a priori* knowledge of the system's behaviour as possible; otherwise spurious results may easily be obtained. Hence the analysis is not conducted as if the system were a completely unknown black box. Rather, certain basic cause-effect relationships are assumed and these assumptions are made by reference to knowledge of the physical, chemical, and biological phenomena that are believed to govern the system's behaviour. The important point is that the use of alternative model forms and alternative algorithms for parameter estimation may yield different insights into the relationships underlying the observed patterns in the field data (see, for example, Beck, 1978).

*Identification of water quality models*

A water quality model, even a relatively simple one, will contain a number of parameters which have to be estimated from the data; it may otherwise be possible to obtain values for the parameters from the literature or from laboratory experiments. The problem of parameter estimation is discussed extensively in part I of this paper (Whitehead *et al.*, 1979) in relation to rainfall-runoff modelling. Rather than use the conventional, large hydrological models for the rainfall-runoff process, a time series modelling approach was developed in which parameters of a simple model were estimated directly from rainfall and runoff data.

In the case of water quality models, this direct parameter estimation approach is particularly important because of the additional complexity introduced by chemical and biological interactions. Model structure identification is an essential and important part of any water quality modelling study (Beck, 1979b). It is often the case that decay coefficients obtained from the literature or laboratory experiments are based on poor data and are therefore inaccurate, or are unique to a particular set of local or ambient conditions. Thus, parameters in the model should always be estimated using data collected during carefully planned field studies.

The most popular approach to determining the model parameters is to use some form of "trial and error" fitting procedure in which model parameters are adjusted until the model output provides the best fit to the corresponding field observations. However, given a relatively limited data set and the inevitable errors associated with sampling and laboratory analysis, this problem is essentially statistical in nature. A range of techniques exist within the control and hydrological literature for estimating the parameters of water quality models (Eykhoff, 1974; Beck, 1980) and during the Bedford Ouse Study two different approaches have been considered. The first uses the extended Kalman filter (Jazwinski, 1970), with which Beck and Young (1976) have identified a model of DO–BOD (Dissolved Oxygen–Biochemical Oxygen Demand) in the River Cam. The second approach employs a multivariable extension of the Instrumental Variable estimation technique introduced in part I of this paper and described in detail by Young & Whitehead (1977).

The advantage of both these techniques is that they formalize the "trial and error" fitting procedure and

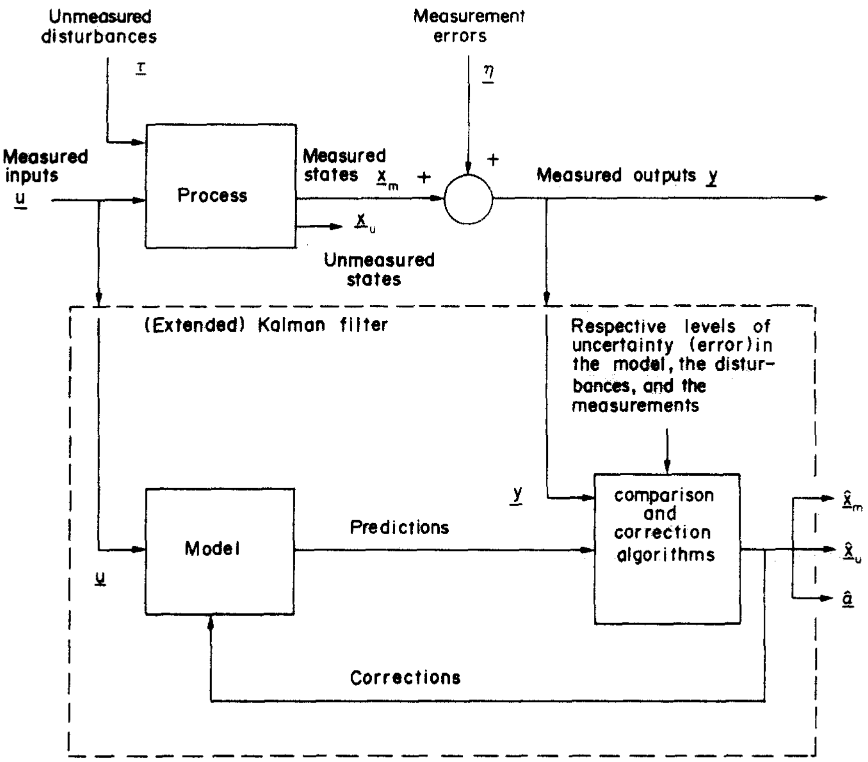


Fig. 4. Conceptual diagram of the extended Kalman filter.

account explicitly for the uncertainties in the *a priori* knowledge and observed behaviour of the system. Another important feature of these techniques is that they provide recursive estimates of the parameters. In other words the model parameters are updated one step at a time moving sequentially through the data. This enables the algorithms to track or follow parameters that vary over time and it provides an additional check on model adequacy. Modellers often conclude, incorrectly, that the model parameters remain constant over the fitting period. The variation of parameters with time is illustrated in part I of this paper in relation to rainfall-runoff modelling and in the case of water quality the utility of the approach has been demonstrated elsewhere (Beck & Young, 1976; Bowles & Grenney, 1978; Whitehead, 1979).

The technique of the extended Kalman filter (EKF), shown schematically in Fig. 4, formalizes the conventional "trial and error" procedure of model calibration. In the "trial and error" approach a deterministic simulation model is run repeatedly through the time-series of data with model parameters adjusted between each run until the predicted behaviour is close to the observed behaviour. Such an approach does not explicitly account for the random disturbances, $\zeta$, which affect river systems or the measurement errors, $\eta$, associated with the observations, y. The EKF, however, accounts for these factors explicitly and provides estimates of the measured state variable (determinands), and the set of parameters, such as decay coefficients, which appear in the model.

A detailed description of the EKF is presented in Jazwinski (1970) and Beck & Young (1976) have described its application to water quality modelling. The approach is very similar to the time series algorithms discussed in part I of this paper and can in fact be shown to be derived from the basic principal of linear regression analysis (see, for example, Young, 1974, Beck, 1979c). The EKF is a recursive algorithm in which an estimate of the unknown parameter vector $\alpha$ is updated while working serially through the data. In this case, for example, the estimate $\hat{\alpha}$ of $\alpha$ at the kth instant in time is given by an algorithm of the general form:

$$\hat{\alpha}_k = \hat{\alpha}_{k-1} + G_{k|k-1}\{y_k - \hat{y}_{k|k-1}\} \tag{5}$$

where the second term on the right hand side is a correction factor based on the difference between the latest determinand measurement $y_k$ and the estimate $\hat{y}_{k|k-1}$ of that determined derived from the model using estimated model coefficients obtained at the previous instant. $G_{k|k-1}$ is a weighting matrix whose elements are calculated essentially as a function of the levels of uncertainty (or error) specified for the model (as an approximation of reality), in the unmeasured input distances ($\zeta$) and in the output response observations ($\eta$). Further details of the interpretation of this matrix can be found in Beck (1979c).

## SAMPLING PROGRAMME AND LABORATORY ANALYSIS

In order to determine rapid changes in river quality a continuous record of water quality determinands is

Table 1. Analysis errors (based on 52 duplicates)

| Determinand | Mean | % Error |
|---|---|---|
| Chloride | 65.10 | 0.94 |
| Ammonia | 3.13 | 1.43 |
| TON | 8.51 | 2.46 |
| BOD | 4.17 | 14.27 |
| DO | 9.28 | 1.13 |
| Sulphate | 164.60 | 6.16 |
| Phosphate | 1.45 | 3.79 |
| Total hardness | 411.50 | 2.40 |
| Alkalinity | 244.70 | 3.49 |
| Non-carbonate hardness | 177.10 | 7.68 |
| Conductivity | 842.60 | 1.82 |
| Suspended solids | 13.05 | 12.29 |
| TOC | 9.42 | 5.83 |

All units $mg\,l^{-1}$ with the exception of conductivity—units mOhm $cm^{-1}$.
* % error = standard deviation × 100/mean.

required. It is not generally sufficient to monitor quality at weekly or even daily intervals, as peak values occurring between samples could not be recorded and information of dynamic behaviour thus lost. In addition, the residence time of some shorter reaches of the river may be of the order of hours and, in these cases, this represents the minimum acceptable sampling interval. However, it was considered at the outset of the Bedford Ouse study in 1972 that the technology of water quality monitoring was such that *only* dissolved oxygen and temperature could be satisfactorily monitored on a continuous basis. It was therefore clearly impossible to measure continuously all the determinands of interest and a daily sampling scheme was proposed together with a series of automatic sampling experiments to investigate the diurnal variations of selected water quality characteristics. This time interval was also considered to be commensurate with the management objectives of the study.

An intensive survey during the summer of 1973 was carried out to assess the feasibility of daily sampling at a number of sites on the river between Newport Pagnell and Clapham (see Fig. 1). The sampling points represented the upstream and downstream boundaries of three reaches along this stretch of the river and daily samples were collected and analysed for a range of water quality variables (Bedford Ouse Study, Final Report, 1979). Less frequent samples were also taken of the larger streams and effluents discharging to the reaches, and, in addition, continuous records of dissolved oxygen (DO) and temperature were obtained using automatic monitors.

The 10 week survey completed over the summer of 1973 demonstrated the feasibility of collecting the volume of data required for a dynamic model and subsequent analysis of these data revealed that acceptable dynamic water quality models could be derived for such variables as DO, Biochemical Oxygen Demand (BOD), Chloride (Cl) and Nitrate (TON). It was suspected, however, that a model derived over an essentially stable low-flow period might not be applicable under different flow regimes or at other times of the year. It was decided therefore to implement a much longer term survey of the river and samples were taken for the period October 1973 to May 1975. The combined data set from the 1973 summer survey and the later prolonged survey provided an extensive record of data with which to establish dynamic stochastic models and develop techniques for the rapid analysis of such data.

In order to assess the laboratory and sampling errors associated with the water quality data, duplicate samples of river water were collected and analysed in the laboratory. Table 1 summarises the percentage errors obtained and indicates the large errors on variables such as BOD. This information is particularly useful during model development since it assists the analyst in specifying some of the statistical properties of the observations used for identification and parameter estimation.

In order to obtain an estimate of any persistent diurnal variation in water quality, samples were obtained at 4-h intervals over a number of days and the data were analysed by computing the mean concentration at each sample time. With the exception of DO and temperature, which are monitored continuously and are known to vary diurnally, the diurnal variations of other determinands were not significant.

Table 2. Average diurnal variations in water quality at Clapham on Bedford Ouse

| Time | 12.30 | 16.30 | 20.30 | 00.30 | 04.30 | 08.30 |
|---|---|---|---|---|---|---|
| Suspended solids | 10.7 | 11.8 | 12.1 | 12.6 | 12.1 | 12.3 |
| S.S. ash | 4.6 | 6.8 | 15.2 | 6.6 | 6.9 | 7.1 |
| Chloride | 52.7 | 51.5 | 50.7 | 50.5 | 51.0 | 50.8 |
| Ammonia | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| TON | 3.9 | 3.9 | 3.9 | 4.0 | 3.9 | 4.1 |
| BOD | 3.5 | 4.3 | 4.1 | 4.2 | 4.1 | 4.1 |
| Sulphate | 136.8 | 131.0 | 121.2 | 134.0 | 122.8 | 136.0 |
| Phosphate | 0.6 | 0.5 | 0.6 | 0.7 | 0.7 | 0.6 |
| Total hardness | 384.7 | 378.7 | 381.3 | 380.7 | 378.0 | 380.7 |
| Alkalinity | 251.7 | 246.7 | 258.7 | 254.2 | 249.2 | 250.8 |
| Non-carbonate hardness | 133.0 | 132.0 | 122.7 | 126.5 | 128.8 | 129.8 |
| Conductivity | 752.0 | 778.0 | 754.0 | 757.5 | 755.0 | 768.0 |
| TOC | 9.9 | 10.0 | 9.5 | 9.3 | 8.7 | 10.9 |

All units $mg\,l^{-1}$ with the exception of conductivity—units mOhm $cm^{-1}$.

For example, the results obtained at Clapham, as shown in Table 2 are typical and the diurnal variations do not appear important, although these may become significant when treated sewage is discharged from Milton Keynes.

## APPLICATION OF THE EKF TO MODEL IDENTIFICATION

The initial analysis of the Bedford Ouse Study data involved the application of the EKF to identify models for chloride, nitrate, DO and BOD using the 1973 summer survey data.

### *Application to modelling conservative substances*

The nature of a conservative determinand is such that a simple mass balance over a river reach should suffice for describing its variations in time. However, the reach is subject to runoff effects and these may be estimated with the EKF technique. The mass balance model for chloride has been assumed to take the form of the first order differential equation, discussed previously, equation (1), i.e.

$$\frac{dx_1(t)}{dt} = \frac{Q(t)}{V}u_1(t) - \frac{Q(t)}{V}x_1(t) + C_r(t) + \zeta_1(t) \quad (6)$$

where $x_1$ is the chloride concentration at the downstream boundary of the reach (mg $l^{-1}$), $u_1$ is the chloride concentration at the upstream boundary of the reach (mg $l^{-1}$), $C_r$ is the net effect of runoff on the in-stream concentration of chloride (mg $l^{-1}$ day$^{-1}$), $Q$ is the streamflow rate (m$^3$ day$^{-1}$) and $V$ is the mean reach volume (m$^3$).

Here, as previously in the study of the River Cam (Beck & Young, 1976), the inclusion of a hypothetical, delayed input ($\tilde{\mathbf{u}}(t)$), see equation (2), did not significantly affect the analysis. The transportation delay component of the idealised reach model has thus been omitted in the following by making the simplifying assumption $\tilde{\mathbf{u}}(t) = \mathbf{u}(t)$ in equation (2); in other words, $\tilde{\mathbf{u}}_1(t) = u_1(t)$ has been substituted in the derivation of equation (6)

The EKF approach to model identification described above was applied to estimate the term $C_1$ using the data obtained during the summer 1973 survey in the reach between the proposed Milton Keynes effluent discharge point and a weir close to Olney (see Fig. 1) some 13 km downstream. As shown in Fig. 5, the runoff term $C_r$ is estimated as being negative over the entire period, suggesting either decay of chloride in the reach (which is unacceptable if chloride is considered conservative) or a dilution effect due to runoff. In this connection, the tributary streams entering the reach were sampled during the summer survey and were consistently lower in chloride than the river; the higher concentrations in the main river are due to the discharge of effluents upstream of Newport Pagnell. The effect of the runoff and tributary stream dilution is particularly noticeable during storm events; e.g. on day 10 a significant localized storm occurred and this

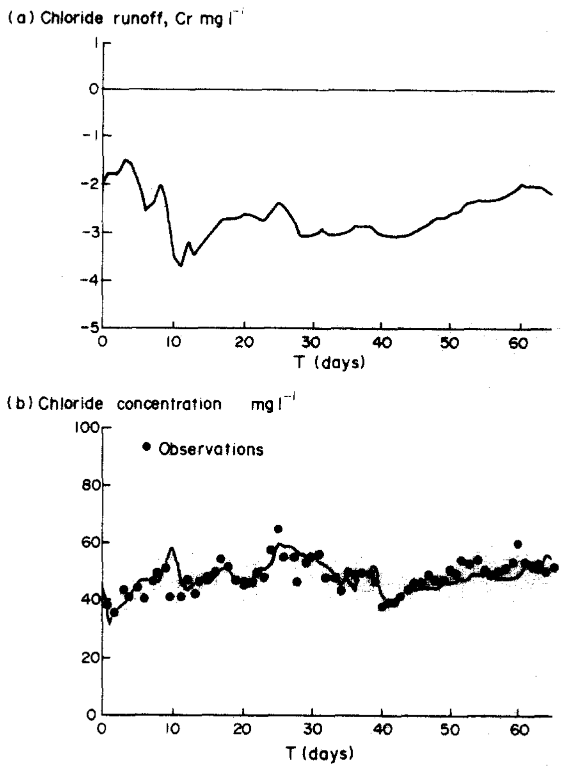


Fig. 5. (a) Estimated chloride runoff. (b) Simulated and observed chloride concentrations.

is shown in Fig. 5 as a sudden change in the estimated runoff effect. When these runoff effects are directly included in the model, the simulated downstream chloride concentrations are close to the observed levels as shown in Fig. 5.

The basic mass balance model thus provides an adequate representation of chloride within the reach and the EKF may be effectively used in an interpretive role to yield information on the runoff effects.

### *Application to modelling non-conservative substences*

The EKF has also been applied to nitrate, DO and BOD models similarly based on the general description of equation (1):

Nitrate

$$\frac{dx_2(t)}{dt} = \frac{Q(t)}{V}u_2(t) - \frac{Q(t)}{V}x_2(t) - k_1x_2(t) + N_R + \zeta_2(t) \quad (7a)$$

BOD

$$\frac{dx_3(t)}{dt} = \frac{Q(t)}{V}u_3(t) - \frac{Q(t)}{V}x_3(t) - k_2x_3(t) + L_A + \zeta_3(t) \quad (7b)$$

DO

$$\frac{dx_4(t)}{dt} = \frac{Q(t)}{V}u_4(t) - \frac{Q(t)}{V}x_4(t) + k_3[C_s(t) - x_4(t)] - k_2x_3(t) + D_B + \zeta_4(t) \quad (7c)$$

where $x_2(t)$ is the downstream nitrate concentration ($mg\,l^{-1}$), $x_3(t)$ is the downstream BOD concentration ($mg\,l^{-1}$), $x_4(t)$ is the downstream DO concentration ($mg\,l^{-1}$), $u_2(t)$, $u_3(t)$ and $u_4(t)$ are the upstream nitrate, BOD and DO concentrations respectively ($mg\,l^{-1}$); $N_R$, $L_A$ and $D_B$ are the lumped effects of runoff and other sources and sinks affecting nitrate, BOD and DO respectively ($mg\,l^{-1}\,day^{-1}$); $C_s(t)$ is the oxygen saturation level ($mg\,l^{-1}$) determined from a polynomial of the form

$$C_s(t) = 14.541233 - 0.3928026\phi(t) + 0.00732326\phi(t)^2 - 0.00006629\phi(t)^3$$

where $\phi(t)$ is the stream temperature (°C). Finally, $k_1$, $k_2$, and $k_3$ are rate constants associated with loss of nitrate, BOD decay, and reaeration respectively ($day^{-1}$).

Using the summer 1973 data the decay parameter associated with nitrate loss, $k_1$ is estimated using the EKF assuming measured values of $N_R$ for the addition of nitrate via runoff. As shown in Fig. 6, the decay rate $k_1$ in the nitrate model is estimated as being reasonably constant at 0.1 over the summer period and the simulated downstream nitrate levels are close to the observed levels. This nitrate loss parameter may be interpreted in terms of the average loss per unit area per day along the reach. Given the average summer flow rate of $3\,m^3s^{-1}$, the average nitrate concentration of $5\,mg\,l^{-1}$ and the mean width of the river bed as 20 m, the average nitrate loss along the reach based on the estimated decay coefficient may be calculated as $0.78\,g\,m^{-2}\,day^{-1}$. This loss compares with the figure of $0.75\,g\,m^{-2}\,day^{-1}$ obtained by Owens *et al.* (1972) under summer conditions from a catchment-wide nutrient budget. The close agreement obtained is surprising since the short reach is not necessarily representative of the entire Bedford Ouse system. However, the fact that the figures are of the same order of magnitude is again a reflection of the effectiveness of the EKF in providing useful information on the system's behaviour.

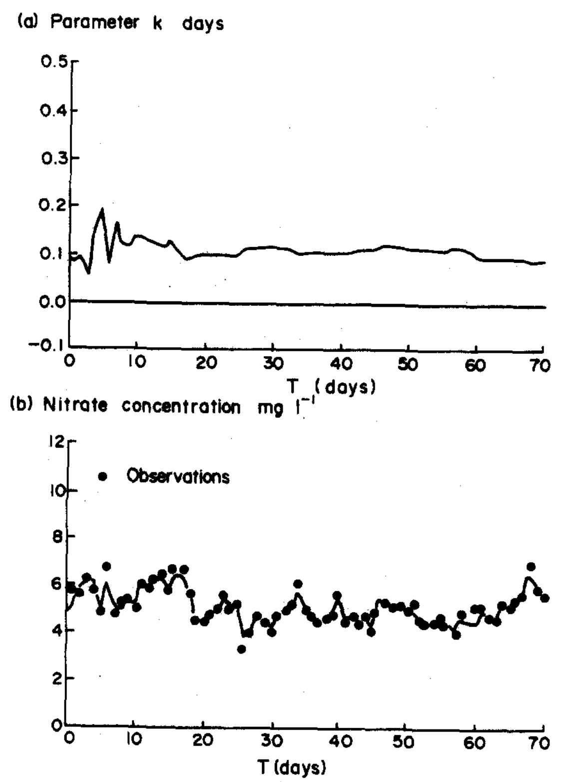


Fig. 6. (a) Estimated nitrate loss parameter, $k_1$. (b) Simulated and observed nitrate concentrations.

In the case of DO and BOD the reaeration and BOD decay constants are estimated in addition to the lumped source and sink terms $D_B$ and $L_A$. Figure 7 shows that for the DO source term, $D_B$ remains positive over the summer period indicating a net input of oxygen. The oxygen production varies according to sunlight conditions and algal population levels and an empirical term to account for these is incorporated into the model (see also Beck & Young, 1975). A modified form of an empirical relationship developed by the Walter Pollution Research Laboratory (now Water Research Centre), 1968, was used, whereby

$$D_B = k_4(0.95 + 0.0317\,\text{Chl-}a)(I - I_t)$$

where $k_4$ is a parameter to be estimated, Chl-$a$ is the chlorophyll concentration ($\mu g\,l^{-1}$), $I$ denotes integrated hours of sunlight (see Beck & Young, 1975) and $I_T$ is a threshold level for the hours of sunlight below which there is no net addition of DO by photosynthesis (i.e. $I - I_T = 0$ for $I < I_T$).

The addition of BOD by mass death algae is described by a relationship utilised by Beck & Young (1975) where

$$L_A = k_5\,(I - I_T)$$

where $k_5$ is a parameter to be estimated.

Applying the EKF in order to estimate the parameters $k_2$, $k_3$, $k_4$ and $k_5$ gave final values of 0.3, 0.35, 0.2 and 0.09 respectively. As indicated in Fig. 7 the parameters are reasonably constant over the data period and, as shown in Fig. 8, the model simulations are satisfactory. Recently, the EKF technique has been applied to modelling variations of ammonia in the Bedford Ouse where it was necessary to estimate nitrification rate constants and details of this application are given elsewhere (Whitehead, 1980b). All of these results indicate that dynamic modelling of water quality in the Bedford Ouse is feasible and that the extended Kalman filter technique may be applied to the problems of model structure identification.

The models developed from the summer survey data provide a basis for the analysis of the data collected from October 1973 to May 1975. In addition to the differential equation models, the alternative discrete time models of water quality discussed before have been developed and these are described in detail elsewhere (Young & Whitehead, 1977). Typical simulations for nitrate, for example are shown in Fig. 9 for

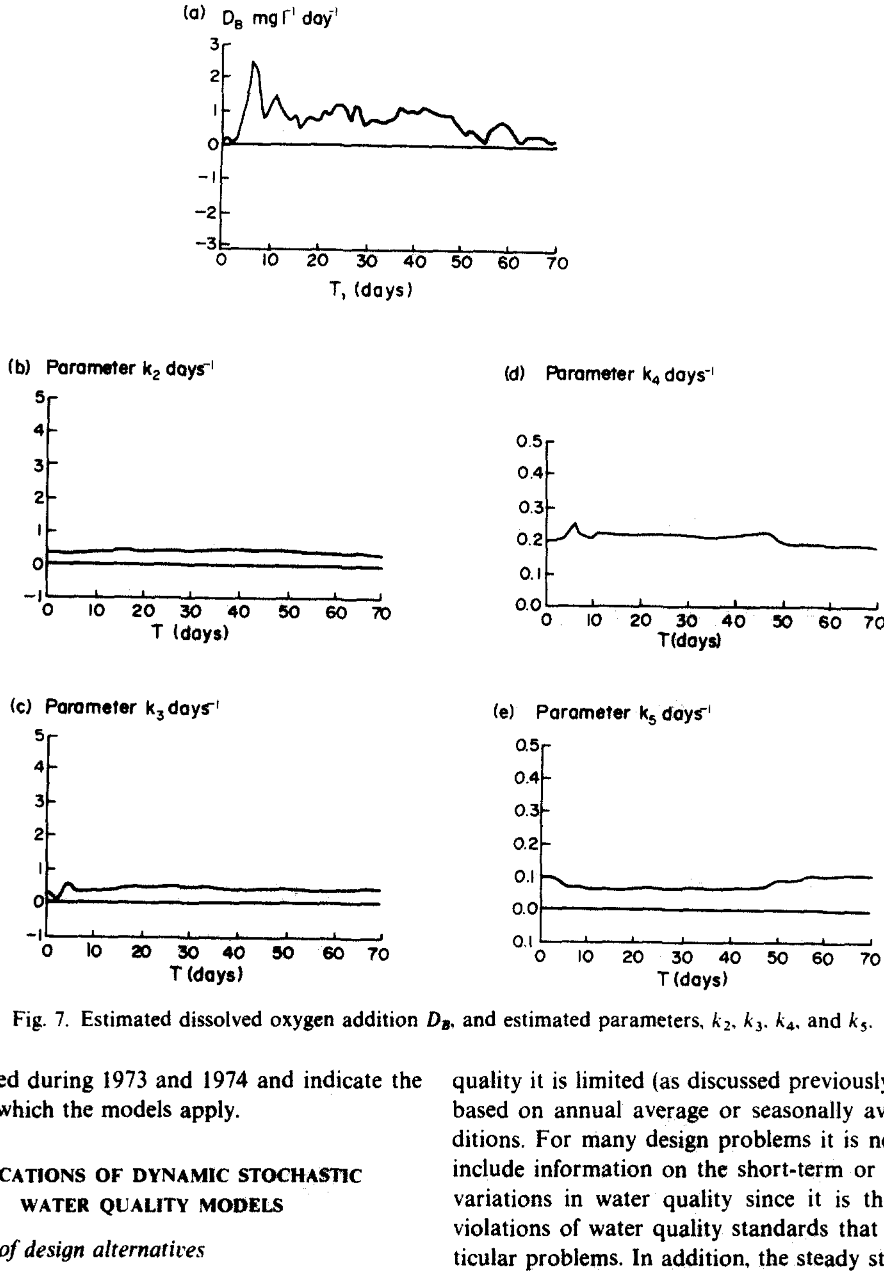


Fig. 7. Estimated dissolved oxygen addition $D_B$, and estimated parameters, $k_2$, $k_3$, $k_4$, and $k_5$.

data collected during 1973 and 1974 and indicate the range over which the models apply.

## APPLICATIONS OF DYNAMIC STOCHASTIC WATER QUALITY MODELS

### *Assessment of design alternatives*

The objectives of the Bedford Ouse Study, as discussed in the introduction to this paper, were to investigate design and operational problems in the Bedford Ouse basin.

A major area of application of water quality models is in the assessment of design alternatives. Because of the high capital costs associated with new structures or facilities for water quality control it is necessary to have a methodology for comparing alternative designs. Models relate cause and effect in water quality and if extended to include cost data, can be used to assess the likely costs and benefits of alternative management policies. During the Trent study (Newsome *et al.*, 1971) and the Bedford Ouse Study (Fawcett, 1975), steady state models were developed for planning purposes. Whilst the steady state model provides a guide to the long term changes in water quality it is limited (as discussed previously) by being based on annual average or seasonally average conditions. For many design problems it is necessary to include information on the short-term or day-to-day variations in water quality since it is the transient violations of water quality standards that cause particular problems. In addition, the steady state models developed to date do not account for the uncertainties associated with sampling, laboratory measurement and imprecise knowledge of the system mechanisms.

At the more detailed design level dynamic stochastic models can be effectively used to assess the impact of effluent on the river system and to design appropriate standards for effluent treatment. For example, the differential equation models of water quality are based on mass balance principles and, therefore, the likely concentration of determinands downstream of a discharge may be determined by simulation. Figure 9 shows the effect on downstream nitrate levels assuming an effluent flow from Milton Keynes of $114{,}000\,m^3\,day^{-1}$ with nitrate levels of $10\,mg\,l^{-1}$. During high-flow water conditions the impact of the effluent is minimal because of dilution effects, and upstream sources of nitrogen and runoff effects pre-

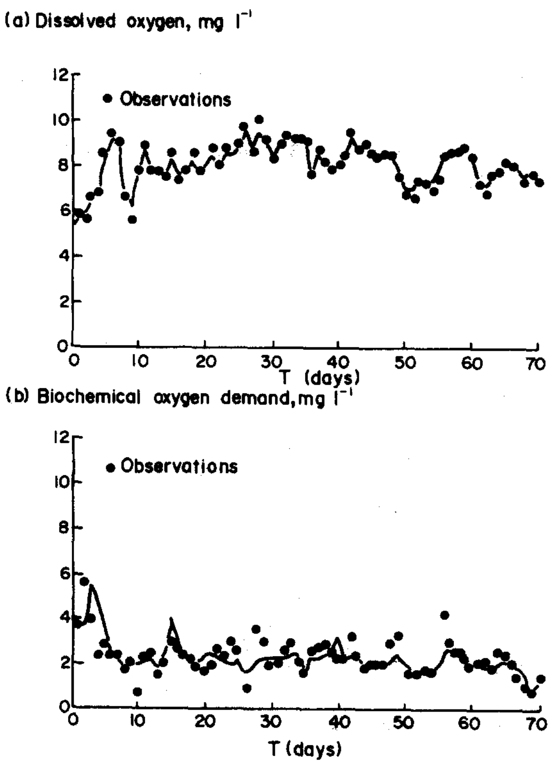


Fig. 8. Simulated and observed DO and BOD.

dominate. In this situation nitrate treatment at Milton Keynes would have relatively little effect and alternative methods of overcoming the high nitrate levels are required such as blending with groundwater or reservoir water at the abstraction plant at Bedford. During low flow conditions and with increased temperature levels during summer, the background levels of nitrogen fall, and the effluent effect becomes more significant.

In addition to providing time varying concentrations at the downstream reference point, the models may be used in a Monte Carlo simulation study to provide forecasts in terms of probability distributions rather than as unique point values. Such a stochastic simulation approach is extremely useful where analytical solutions are difficult or even impossible to obtain, as is often the case with reasonably complicated dynamic systems. The calculations (usually simulations) are performed a large number of times, each time with the values for the stochastic inputs or uncertain parameters selected at random from their estimated parent probability distributions (see Fig. 10). Each such random experiment or simulation yields a different result for any variable of interest and when all these results are taken together the required probability distribution can be ascertained to any degree of accuracy from the sample statistics. The degree of accuracy of the probability distribution function estimated in this manner is, of course, a function of the number of random simulations used to calculate the sample statistics, but it is possible to quantify the degree of uncertainty on the distribution using non-parametric statistical tests such as the Kolmogorov–Reyni statistics (see Spear, 1970).

Monte Carlo simulation is a flexible, albeit computationally expensive tool with which to investigate certain design problems. For example, the water quality standards proposed in the Bedford Ouse Study (final report, 1979) are presented in terms of probability distribution functions, and therefore pro-

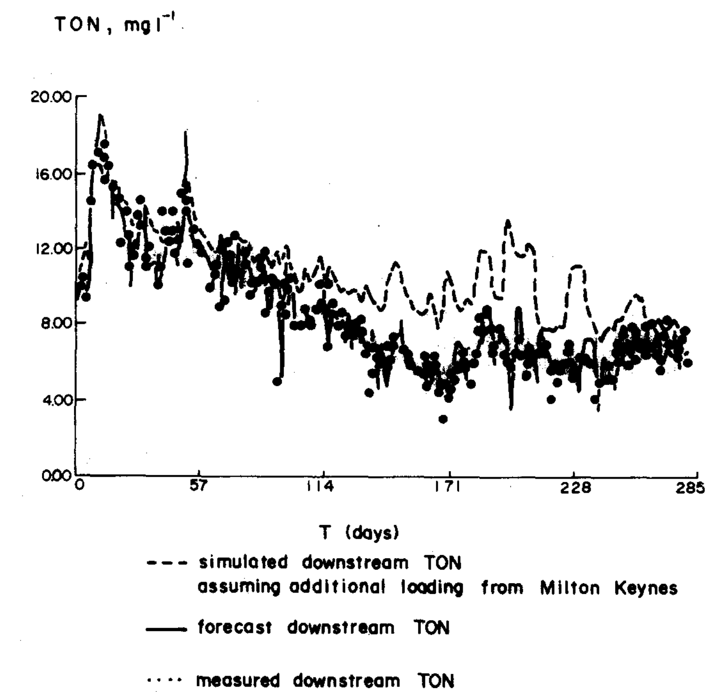


Fig. 9. Nitrate simulations and observations on the Bedford Ouse during 1974.

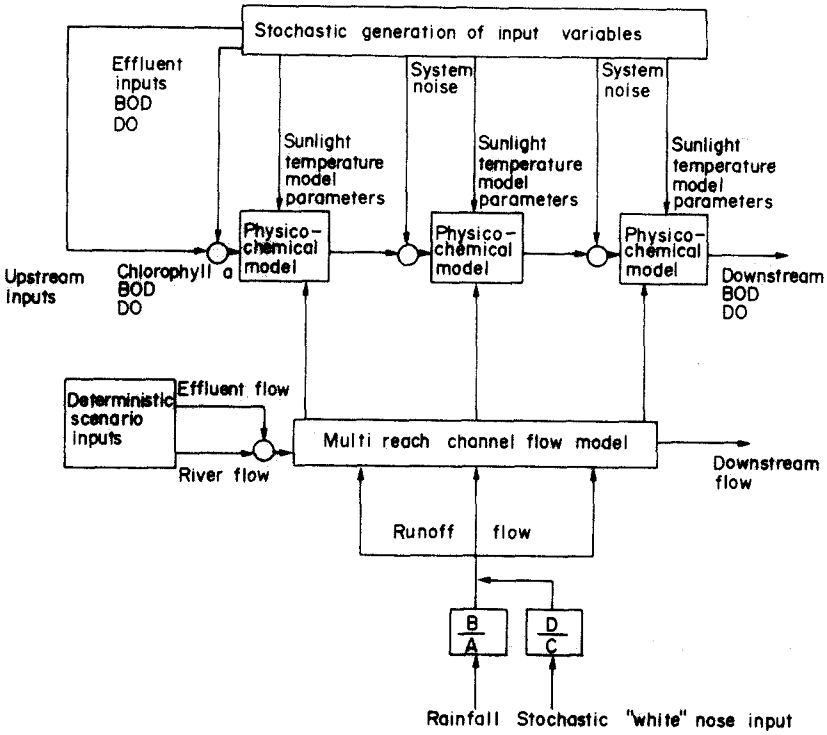


Fig. 10. Monte Carlo simulation model.

vide a reference against which the water quality can be tested. It would be possible to perform Monte Carlo simulation analysis using the water quality models developed for the study section of the Bedford Ouse under various assumptions about future levels of effluent input. The outcome of such an analysis would be probability density functions for the water quality states which could be compared directly with the required water quality standards. Such information would be extremely useful in assessing the impact of influent on the system and in determining the degree of treatment necessary at Milton Keynes in order to ensure satisfactory water quality at the abstraction point.

An *initial* assessment of the impact of Milton Keynes effluent on the aquatic environment has been obtained using Monte Carlo simulation, details of which are given by Whitehead & Young (1979). Altogeather three effluent conditions were considered at different flow rates and BOD levels, as shown in Table 3.

Table 3.

| Effluent rate ($m^3 s^{-1}$) | BOD concentration in effluent ($mg\,l^{-1}$) | Variance of BOD concentrations |
|---|---|---|
| Case 1 0.1 | 5 | 1 |
| Case 2 0.4 | 10 | 4 |
| Case 3 1.0 | 10 | 4 |

It was assumed that the effluent has no dissolved oxygen present; this is clearly a condition that represents the worst situation but which is not unrealistic as the effluent is pumped direct from the treatment works via a 4-km pipe into the river. Effluent BOD levels fluctuate in practice and a stochastic component defined by a noise signal of known variance (see Table 3) was added to the three BOD levels. The distributions of BOD and DO at Bedford given these three effluent conditions are compared with present situations in Fig. 11. At low discharge conditions there is relatively little effect on the aquatic environment. At the $1\ m^3 s^{-1}$ condition, however, the mean BOD level has risen to $4.5\ mg\,l^{-1}$ and mean DO level has fallen to $6.5\ mg\,l^{-1}$ with the DO distribution ranging from 4.5 to $9\ mg\,l^{-1}$. These distributions represent only an initial assessment of the impact of Milton Keynes effluent and, in fact, the DO levels may be adversely affected by the changing biological nature of the river.

*Operational aspects of water quality management*

The second area of application of the Bedford Ouse dynamic models lies in the management and operational control of water resource systems (Whitehead, 1978). A major requirement of operational managers is information on the present condition of the river system and on likely future changes in water quality. Operational managers must be able to respond quickly to emergency situations in order to protect and conserve the river and maintain adequate water supplies for public use. Moreover, the costs of water

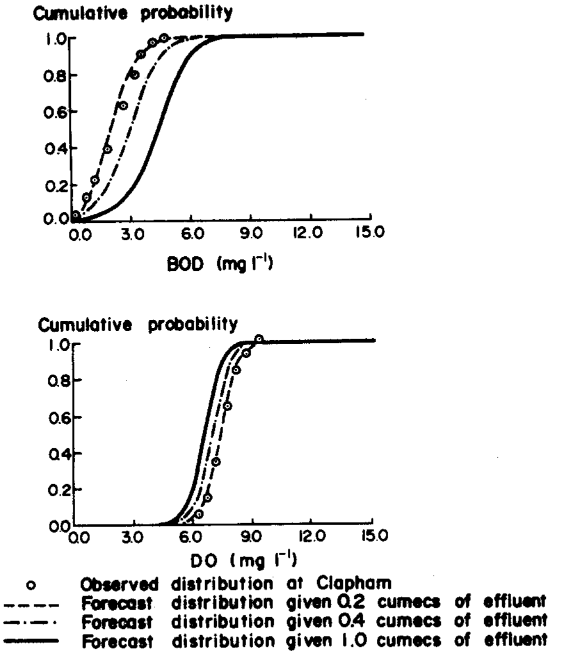


Fig. 11. Cumulative probability distributions for DO and BOD assuming different effluent loadings.

treatment and bankside storage are particularly high and there are therefore considerable benefits to be gained from the efficient operational management of river systems from the viewpoint of water quality.

In recent years there has been some progress towards providing more efficient operational management by the installation of automatic continuous water quality monitors on river systems. These measure such water quality variables as dissolved oxygen, ammonia, and temperature and, if combined with a telemetry scheme relaying information to a central location, provide immediate information on the state of the river for pollution officers. Whilst the reliability of such schemes is still rather poor, there is now an opportunity to use this information together with mathematical models for making real-time forecasts of water quality. Indeed mathematical models can be used to improve reliability by detecting drift or systematic errors in monitors at an early stage and by interpolating to allow for missing data points.

The practical problems associated with continuous field measurement and telemetry of water quality variations have largely limited the application of on-line forecasting and control schemes. The continuous flow of water past sensors for measuring water quality gives rise to severe fouling of optical and membrane surfaces, thereby drastically reducing the accuracy of the data produced. In recent years, however, there have been several studies and applications of continuous water quality monitors (Briggs, 1975; Kohonen *et al.*, 1978). Most U.K. Water Authorities have established monitoring and telemetry schemes (Hinge & Stott, 1975; Caddy & Akielan, 1978) and report reasonable reliability provided the monitors are regularly maintained. More recently, Wallwork (1979) has described an application on the River Wear in North East England where a continuous monitor is used to protect an abstraction point.

The application of particular interest here is an extensive monitoring and telemetry scheme which has been developed since 1975 along the Bedford Ouse River system in South Eastern England.

As indicated in Fig. 12 automatic water quality monitors have been installed at several sites along the river and data on such variables as dissolved oxygen, conductivity, nitrate, ammonia, and temperature are telemetered at frequent intervals to the central control station located in Cambridge. The telemetry also transmits information on flow at key gauging stations and a micro-computer located in Cambridge is used to analyse the date on-line. The system provides rapid information on the present state of the river and incorporates the dynamic water quality model for making real-time forecasts of flow and quality at key locations along the river system. The models are applied, however, using a much shorter time unit than a day (e.g. 1 h) to provide up to date information for management). The functions of the microcomputer may be summarized as:

(a) scanning the automatic water quality monitoring stations and flow gauges at frequent intervals;

(b) transforming the data to relevant physical and chemical parameters (eg river level to flow via stage-discharge equations or converting nitrate selective ion electrode measurements to concentration units);

(c) data reduction (such as the calculation of daily means, maxima and minima), data storage and printing of data summaries for management;

(d) equipment calibration and warning of failure or drift;

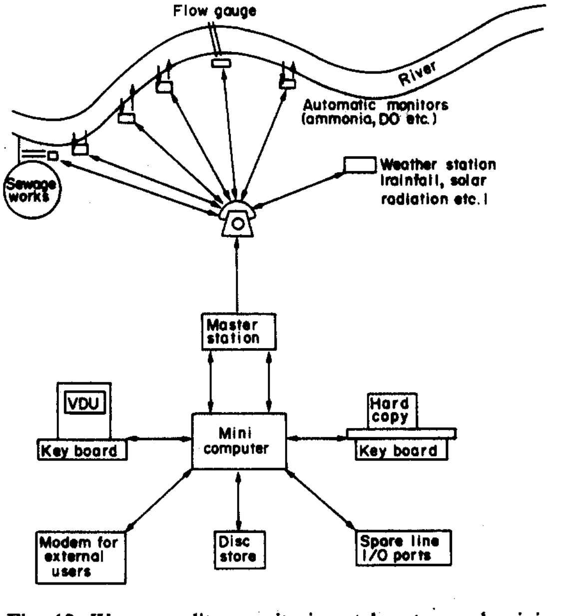


Fig. 12. Water quality monitoring, telemetry and mini-computer processing system for the Bedford Ouse.

(e) alarm signalling in the event of extreme water quality conditions;

(f) forecasting flow and water quality at key locations and the calculation of travel times and dispersion information.

The automatic monitoring scheme on the Bedford Ouse provides real time information on the quality of the river thus enabling pollution officers to detect pollutions at an early stage. With flow and quality models of the system it is possible to predict the time of travel of a pollutant load and the dilution factors along a particular stretch of river and, for example, in the case of Bedford Water Division abstraction plant, provide several hours warning of a pollution event. This information is particularly useful since alternative sources of water such as groundwater or reservoir water are often required for blending purposes and increases levels of treatment such as chlorination or nitrate removal (a future possibility at the Bedford Water Division treatment plant) are necessary to maintain an adequate supply of drinking water.

## CONCLUSIONS

In this paper we have outlined the major aspects of the study of short term water quality variations in a non-tidal river system. As in the case of the flow model developed in part I, we have been able to develop dynamic-stochastic models for the 55 km stretch of the Bedford Ouse between the site of Milton Keynes and the Bedford Water Division abstraction plant, near Bedford. We believe that this model is one of the first examples of an integrated dynamic-stochastic model of flow and water quality for a long stretch of river to be satisfactorily identified and estimated by reference to daily field data.

The research has demonstrated the feasibility of constructing realistic dynamic water quality models for non-tidal river systems: in particular the models described in the paper provide a potentially useful and computationally efficient characterization of the flow and water quality phenomena in a river; a characterization which explains the dynamic behaviour of chloride, nitrate , DO and BOD and so provides a firm foundation on which to base the investigation of other water quality states.

In developing dynamic-stochastic models of water quality it is also felt that we have been able to demonstrate the particular utility of time series analysis techniques for the identification and estimation of water resource models. The techniques have been applied to a wide range of models from the purely black-box rainfall-runoff models, multivariable "grey box" models of water quality, to the more detailed differential equation models of water quality. The EKF and Instrumental Variable methods of recursive estimation provide a powerful general method of data processing well suited to the sort of modelling problems encountered by hydrologists, water quality and water resource engineers.

The models developed during the study are prescriptive and are, therefore, designed to be used by management. The application of the models to both design and operational control and management problems represents an important aspect of the research. The design of a water resource system from the viewpoint of water quality has conventionally been based on "steady state" models which provide information on annual average conditions. However, for many design problems, detailed information on the transient behaviour of water quality is required with a description of the stochastic aspects of water quality. Such information can only be realistically obtained using dynamic stochastic models of flow and water quality of the type developed for the Bedford Ouse. In addition, the technique of Monte Carlo simulation may be effectively used to provide forecasts of water quality directly in terms of probability distributions.

In recent years continuous water quality monitoring schemes have been developed in conjunction with telemetry systems to provide real-time information for operational management. The rapid development in microcomputers have enhanced such schemes by providing considerable analytical power for on-line data processing at relatively low cost. The real time forecasting and control of water quality along critical stretches of river systems is therefore an option now available to operational management. Such an application has been considered for the Bedford Ouse river system and has been implemented by the Anglian Water Authority and the Institute of Hydrology.

*Acknowledgements*—The views expressed in this paper do not necessarily reflect the views of the Anglian Water Authority but the opinions of the authors. The research progress has been largely due to the excellent communications with Dr D Caddy of the Anglian Water Authority and the guidance of Professor P. C. Young. We are also indebted to the many members of the Steering Group of the Associated Committee for the Great Ouse Area and, in particular, N. Taylor, A. Fawcett, D. Newsome, J. Cole, B. Price and M. Owens. Finally we would like to thank P. Garnett and A. Simpson of DOE for their support.

## REFERENCES

Beck M. B. (1978) Random sigma malysis in an environmental sciences problem. *Appl. Math. Model.* **2**, 23–29.

Beck M. B. (1979a) Time variable water quality management—a policy study, working paper WP-79-125. International Institute for Applied Systems Analysis, Laxemburg, Austria.

Beck M. B. (1979b) Model structure identification from experimental data. In *Theoretical Systems Ecology: Advances and Case Studies*, pp. 259–289. Academic Press, New York.

Beck M. B. (1979c) System identification, estimation and forecasting of water quality—Part I: theory working paper, WP-73-31. International Institute for Applied Systems Analysis, Laxemburg, Austria.

Beck M. B. (1980) Applications of system identification and parameter estimation in water quality modelling *Proceedings of the IAHS Hydrological Forecasting Symposium*, pp. 123–131. IAHS Publication No. 129, Oxford.

Beck M. B. & Young P. C. (1975) A dynamic model for DO-BOD relationships in a non-tidal stream. *Water Res.* **9**, 769–776.

Beck M. B. & Young P. C. (1976) Systematic identification of DO-BOD model structure. *Am. Soc. civ. Engrs envir. Engng Div.* **102**, 909–924.

Bedford Ouse Study, Final Report (1979) Anglian Water Authority.

Bowles D. S. & Grenney W. J. (1978) Steady state river quality modelling by sequential extended Kalman filters. *Wat. Resour. Res.* **14**, 85–95.

Briggs R. (1975) Instrumentation for monitoring water quality. *J.S.W.T.E.* **24**, 23–45.

Caddy D. E. & Akielan A. W. (1978) Management of river water quality. *Int. envir. Safety.* **1**, 18–26.

Casapieri P., Fox T. M., Owers P. J. & Thompson G. D. (1978) A mathematical deterministic river quality model. Part 2 Use in evaluating the water quality management of the Blackwater catchment. *Water Res.* **12**, 1155.

Dorf R. C. (1975) Time-domain analysis and design of control systems. Addison-Wesley, Reading, MA.

Eykhoff P. (1974) System identification—parameter and state estimation. Wiley, London.

Fawcett A. (1975) A management model for river water quality, Bedford Ouse Study conference proceedings Anglian Water Authority.

Himmelblau D. M. & Bischoff K. B. (1968) *Process Analysis and Simulation.* Wiley, New York.

Hinge D. C. & Stott D. A. (1975) Experience in the continuous monitoring of river water quality. Instruments and Control Systems Conference, Water Research Centre.

Jamieson D. G. (1978) A hierarchical approach to water quality modelling *Mathematical Models for Water Pollution Control.* Wiley, New York.

Jazwinski A. H. (1970) *Stochastic Processes and Filtering Theory.* Academic Press, New York.

Kohonen T., Hell P., Muhonen J. & Vuolas E. (1978) Automatic water quality monitoring systems in Finland. Report of the National Board of Waters, Finland, No. 153.

Knowles G. & Wakeford A. C. (1978) A mathematical deterministic river quality model. Part 1: formulation and description. *Water Res.* **12**, 1149.

Newsome D. H., Bowden K. & Green J. A. (1971) Trent mathematical model. *Symposium on Advanced Techniques on River Basin Management, the Trent Model Research Programme*, Institution of Water Engineers.

O'Neil P. G. (1972) A mathematical programming model for manning a regional water resources system. *J. Inst. Water Engrs* **26**,

Onstad. C. A. & Blake J. (1980) Thames nitrate and agricultural relations. *Proceedings of ASCE Watershead Management Symposium*, pp. 961–973. Boise, I.D.

Owens M., Garland J., Hart I. C. & Wood G. (1972) Nutrient budgets in rivers. *Symp. Zool. Soc. Land.*, No 29, pp. 21–40.

Page C & Warn A. E. (1974) Water quality considerations in the design of water resource systems. *Water Res.* **8**, 969–975.

Rinaldi S., Soncini-Sessa R., Stehfest H. & Tanuara H. (1979) *Modelling and Control of River Quality.* McGraw-Hill, New York.

Spear R. C. (1970) The application of Kolmogorov–Renyi statistics to problems of parameter uncertainty in systems design. *Int. J. Control.* **11**, 771–778.

Stott P. F. (1979) Water resources in England and Wales—recent experience and current strategy. *Proceedings of Inst. of Civil Engineers Conference*, London.

Thomann R. V. (1972) *Systems Analysis and Water Quality Management.* McGraw-Hill, New York.

Thompson G. D. (1979) A model for nitrate-nitrogen transport and denitrification in the R. Thames. *Water Res.* **13**, 855.

Wallwork J. F. (1979) Protecting a water supply intake—river water data collection and pollution monitoring. River Pollution Control Conference, Water Research Centre.

Warn A. E. (1980) Water Quality Mass Balance Models. *Inst. Wat. Pollut. Control J.* To be published.

Water Research Centre Annual Report (1968).

Whitehead P. G. (1978) Modelling and operational control of water quality in river systems. *Water Res.* **12**, 377–384.

Whitehead P. G. (1979) Applications of recursive estimation techniques to time variable hydrological systems. *J. Hydrol.* **40**, 1–16.

Whitehead P. G. (1980a) An instrumental variable method of estimating differential equations of dispersion and water quality in non tidal rivers. *J. Ecol. Model.* **9**, 1–14.

Whitehead P. G. (1980b) Real time monitoring and forecasting of water quality in the Bedford Ouse river system. *Proceedings of IAHS Conference*, Publication No. 129, Oxford.

Whitehead P. G. & Young P. C. (1979) Water quality in river systems: Monte Carlo Analysis. *Water Resour. Res.* **15**, 451–459.

Whitehead P. G., Young P. C. & Hornberger G. (1979) A systems model of flow and water quality in the Bedford Ouse river system—I. Streamflow modelling. *Water Res.* **13**, 1155–1169.

Young P. C. (1974) A recursive approach to time series analysis. *Bull. Inst. Appl.* **10**, 209–224.

Young P. C. & Whitehead P. G. (1977) A recursive approach to time-series analysis for multivariable systems. *Int. J. Control.* **25**, 457–482.